PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

AD

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [4] : C07H 21/04, C12N 15/00, 9/90 C12N 1/20, G01N 33/564

A1

(11) International Publication Number: WO 89/09222

(43) International Publication Date: 5 October 1989 (05.10.89)

(21) International Application Number: PCT/US89/01116

(22) International Filing Date: 22 March 1989 (22.03.89)

(31) Priority Application Number: 172,159

(32) Priority Date: 23 March 1988 (23.03.88)

(33) Priority Country: US

(71) Applicant: THE JOHNS HOPKINS UNIVERSITY [US/US]; 720 Rutland Avenue, Baltimore, MD 21205 (US).

(72) Inventors: EARNSHAW, William, C. ; 2010 Corbett Road, Monkton, MD 21111 (US). D'ARPA, Peter ; 651 Stirling Street, Baltimore, MD 21202 (US).

(74) Agents: HOSCHEIT, Dale, H. et al.; Banner, Birch, McKie & Beckett, One Thomas Circle, N.W., Sixth Floor, Washington, DC 20005 (US).

(81) Designated States: AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CF (OAPI patent), CG (OAPI patent), CH, CH (European patent), CM (OAPI patent), DE, DE (European patent), DK, FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL, NL (European patent), NO, RO, SD, SE, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent).

Published
With international search report.

(54) Title: CLONED HUMAN TOPOISOMERASE I

(57) Abstract

The invention relates to cDNA molecules coding for eukaryotic topoisomerase I polypeptide. The invention provides cDNA molecules comprising a part of the cDNA sequence of topoisomerase I which encode at least one epitope for autoantibodies to eukaryotic topoisomerase I. The invention also provides cloning vehicles capable of replication and expression comprising cDNA molecules coding for eukaryotic topoisomerase I. The invention further provides for hosts transformed with a vehicle having a cDNA molecule coding for eukaryotic topoisomerase I. In another embodiment, the invention provides for the detection of autoantibodies to eukaryotic topoisomerase I using the eukaryotic topoisomerase I polypeptides coded for by the cDNA molecules of the invention.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	ML	Mali
AU	Australia	GA	Gabon	MR	Mauritania
BB	Barbados	GB	United Kingdom	MW	Malawi
BE	Belgium	HU	Hungary	NL	Netherlands
BG	Bulgaria	IT	Italy	NO	Norway
BJ	Benin	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
CH	Switzerland	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
DE	Germany, Federal Republic of	LU	Luxembourg	TG	Togo
DK	Denmark	MC	Monaco	US	United States of America

CLONED HUMAN TOPOISOMERASE I

BACKGROUND OF THE INVENTION

This work was supported a Grant from the National Institutes of Health. The United States Government may retain certain rights in this invention.

FIELD OF THE INVENTION

The present invention relates to the use of recombinant DNA technology for the transformation of a host organism with eukaryotic genetic sequences for the expression of topoisomerase I polypeptides and methods of using topoisomerase I polypeptides to detect auto-antibodies to topoisomerase I polypeptides.

DESCRIPTION OF THE BACKGROUND ART

DNA topoisomerases I and II catalyze the breaking and rejoining of DNA strands in a way that allows the strands to pass through one another, thus altering the topology of DNA. Type I topoisomerases break a single DNA strand, while the type II enzymes break two strands of duplex DNA. Both enzymes can perform a variety of similar topological interconversions, including relaxation of super coiled DNA, knotting/unknotting and catenation/decatenation of duplex DNA.

Both topoisomerases I and II can provide the topological interconversions necessary for transcription and replication. For example, topoisomerase I can provide the necessary unlinking activity for efficient in vitro DNA replication (Minden, et al., Journal of Biological Chemistry, 260:9316, 1985; Nagata, et al., Proceedings of the National Academy of Sciences, U.S.A., 80:4266, 1983; Yang, et al., Proceedings of the National Academy of Sciences, U.S.A., 84:950, 1987) however, topoisomerase II can also facilitate the replication of

SV40 DNA by HeLa cell lysates (Yang, et al., Proceedings of the National Academy of Sciences, U.S.A., 84:950, 1987). Genetic studies in yeast reveal that both replication and transcription proceed in single mutants deficient in either topoisomerase I or II (Goto, et al. Proceedings of the National Academy of Sciences, U.S.A., 82:7178, 1985; Uemura, et al., EMBO Journal, 3:1737, 1984; Thrash, et al., Proceedings of the National Academy of Sciences, U.S.A., 82:4274, 1985). In cells lacking both topoisomerases, transcription and replication are dramatically reduced (Uemura, et al., EMBO Journal, 5:1003, 1986; Brill, et al., Nature, 326:414, 1987).

Several lines of evidence suggest that topoisomerase I normally functions during transcription. The enzyme has been shown to be localized preferentially to actively transcribed loci by immunofluorescence (Fleishmann, et al., Proceedings of the National Academy of Sciences, U.S.A., 81:6958, 1984), and by co-immunoprecipitation with transcribed DNA (Gilmoure, et al., Cell, 44:401, 1986; Muller, et al., EMBO Journal, 1237, 1985). Furthermore, topoisomerase I cleavage sites have been mapped to regions in and around transcribed DNA (Bonner, et al., Cell, 41:541, 1985; Gilmour, et al., Molecular Cell Biology, 7:141, 1987; Stewart, et al., Cell, 50:109, 1987). Nonetheless, at least in yeast, topoisomerase II can apparently substitute for the functions of topoisomerase I in transcription (Uemura, et al., EMBO Journal, 3:1737, 1984; Thrash, et al., ibid).

Of further interest is the use of topoisomerase I in classifying autoimmune disease. Autoimmune diseases are diseases in which an animal's immune system attacks its own tissues. Often the various types of autoimmune disease can be characterized based upon the specificity of autoantibodies which are produced. For example, it is well known that the serum of patients having the connective tissue autoimmune disease progressive systemic sclerosis (PSS), also known as scleroderma, frequently contain antibodies to such nuclear antigens as topoisomerase I. Thus, the ability to accurately detect the presence of antibodies reactive with topoisomerase I can greatly assist in evaluating the prognosis and planning, or monitoring, the appropriate therapy of patients with scleroderma.

Unfortunately, the existing commercial dectection systems utilize crude exacts of cell nuclei, usually of rabbit or bovine origin, which contain antigens other than topoisomerase I. As a result, these systems will detect antibodies which react with nuclear antigens other than topoisomerase I such as, for example, the centromere, and, as a result, may give false positive results which can lead to an incorrect diagnosis. Thus, there is considerable need for a system which detects only antibodies to topoisomerase I and not other eukaryotic peptides.

SUMMARY OF THE INVENTION

It is a major object of the present invention to provide a DNA sequence which encodes eukaryotic topoisomerase I polypeptide.

Another object of the present invention is to provide a prokaryotic or eukaryotic host cell transformed or transfected with a DNA sequence which encodes eukaryotic topoisomerase I polypeptide such that the host cell can express this polypeptide.

Yet another object of the present invention is to provide a biologically functional plasmid or viral DNA vector which includes a DNA sequence encoding eukaryotic topoisomerase I polypeptide.

It is still another object of the present invention to provide a DNA sequence consisting essentially of a DNA sequence encoding a polypeptide which has an amino acid sequence possessing at least one epitope for binding topoisomerase I autoantibodies.

It is an object of the invention to provide a method of producing eukaryotic topoisomerase I polypeptide in host cells.

Another object of the present invention is to provide a method of detecting autoantibodies to eukaryotic topoisomerase I.

These and other objects are provided by one or more of the following embodiments of the present invention. In one embodiment a DNA sequence is provided which encodes eukaryotic topoisomerase I polypeptide.

In another embodiment a biologically functional plasmid or viral DNA vector is provided which includes a DNA sequence encoding for eukaryotic topoisomerase I polypeptide.

In another embodiment, a method of producing eukaryotic topoisomerase I polypeptide is presented which comprises providing a host cell which can replicate and express an intron-free DNA sequence of eukaryotic topoisomerase I polypeptide, growing the host cell, and recovering the eukaryotic topoisomerase I polypeptide. Also contemplated by the present invention are the transformed or transfected host cells themselves.

In yet another embodiment eukaryotic topoisomerase I polypeptide is provided which is substantially free of other eukaryotic non-topoisomerase I polypeptides.

In still another embodiment, a method of detecting autoantibodies to eukaryotic topoisomerase I polypeptide is provided which comprises contacting a sample with the polypeptide expressed by the DNA sequence of the invention, incubating the sample with the polypeptide for a period of time and under conditions sufficient for the autoantibodies to bind to the polypeptide, separating the autoantibodies which are bound to the polypeptide from the rest of the sample, and detecting the presence of the autoantibodies which bind to the polypeptide.

A major advantage of the present invention is that it provides the art with a ready source of eukaryotic topoisomerase I polypeptide corresponding to that purified from natural sources, while avoiding the problems associated with the isolation of naturally occurring eukaryotic topoisomerase I polypeptide to separate it from other eukaryotic non-topoisomerase I polypeptides. This absence of other eukaryotic non-topoisomerase I polypeptides is significant in that it allows the development of test systems which will only detect antibodies specifically reactive with topoisomerase I polypeptides.

Another advantage of providing eukaryotic topoisomerase I polypeptide in host cells is that by so doing it is possible to obtain much larger quantities of the polypeptide than are currently practicably available from natural sources. As a consequence, not only is it possible to use the polypeptide of the invention to more accurately classify patients with such autoimmune rheumatic diseases as scleroderma, but it is also now possible to also provide commercially

useful quantities of topoisomerase I polypeptide for use in diagnostic systems.

The eukaryotic topoisamerase I polypeptides of the present invention retain the ability to bind autoantibodies to eukaryotic topoisomerase I even though the prokaryotic host in which the polypeptides of the invention can be cloned degrades the transcribed topoisomerase I molecules into a spectrum of polypeptides.

DESCRIPTION OF THE DRAWINGS

Figure 1. Restriction map and strategy for sequencing topoisomerase I cDNA clones T1A and T1B. The bold lines represent the cDNA coding sequences and the thin lines the presumptive 5' and 3' untranslated regions. A partial restriction endonuclease map is shown (X, Xho I; M, Xmn I; N, Nde I; H, Hind III; R, Eco RV; E, Eco RI). The arrows represent the individual subclones that were sequenced to determine the complete cDNA sequence. The subclone indicated by the asterisk was used to probe the DNA blot shown in figure 4.

Figure 2. Expression of fp-topo I and comparison of the proteolysis maps of human nuclear topoisomerase I and fp-topo I. A) Immunoblots of lysates of bacteria harboring pATH11 (-) or pATH11-TIA(F) probed with anti-topoisomerase I serum. B) Proteolysis maps. Human topoisomerase I (T) and fp-topo I (F) obtained by excision from $NaDodSO_4$ polyacrylamide gel. Lanes labeled "-" were loaded with gel slices (between 95 and 116kDa) excised from SDS-PAGE gels loaded with bacterial lysates that contained the pATH11 expression vector without the T1A cDNA. V8 protease or chymotrypsin were overlaid so that proteolysis could proceed during subsequent electrophoresis. An immunoblot probed with anti-topoisomerase I autoimmune serum is shown. Molecular weight markers: 200, 116, 95, 68, 60, 48, 40, 29, 18.4, 16.9 kDa.

Figure 3. A) Topoisomerase I activity in lysates of bacteria harboring pATH11 or pATH11-T1A. Lysates of bacteria harboring pATH11 (-) or pATH11-T1A(+) were added to reaction mixtures containing either buffer P (with free Mg++) or buffer E (without free Mg++). B) Immunodepletion of Mg++-independent topoisomerase I activity from bacterial lysates. Lysates of bacteria harboring

pATH11-T1A were immunoprecipitated with either normal human (NH) or with anti-topoisomerase I (alpha -topo) serum. The supernatant (lane 1) and serial two-fold dilutions (lanes 2 and 3) were then assayed for topoisomerase I activity in the absence of free Mg++.

Figure 4. Human topoisomerase I is encoded by a 4.1 kb mRNA transcribed from a single copy human gene. A) Human genomic DNA cut with EcoR I (R) or Hind III (H) was probed with the fragment from T1A which extends from codon 484 to 578 (fragment with asterisk in Figure 1, also see Figure 6). DNA size markers: 23130, 9416, 6557, 4361, 2322, 2027, 1353 bp. B) Polyadenylated mRNA from HeLa cells probed with the entirety of T1A. RNA markers: 9490, 7460, 4400, 2370, 1350 nucleotides.

Figure 5. Sequence of topoisomerase I cDNA. The sequence was performed according to the strategy outlined in Figure 1. The sequence includes the presumptive 5' and 3' untranslated sequences of 211 and 1139 bases, respectively.

Figure 6. Comparison of the amino acid sequences of human (bottom) and Saccharomyces cerevisiae (top) topoisomerase I as deduced from their DNA sequences. A) Double dots represent amino acid identities between S. cerevisiae and human type I topoisomerases. Single dots represent similarity in sequence due to conservative changes. Asterisks denote a large gap placed in the human sequence in order to maintain maximal alignment. B) Schematic diagram of regions of greatest similarity between human and S. cerevisiae amino acid sequences. Thin lines represent regions without apparent similarity while the heavy bars represent regions of significant similarity. Regions shown as heavy bars show 55% identity or 69% similarity when conservative changes are considered.

DETAILED DESCRIPTION

The present invention relates to the manipulation of genetic materials by recombinant procedures which make possible the production of polypeptides possessing part or all of the primary structural conformation for one or more of the epitopes for binding autoantibodies to eukaryotic topoisomerase I. These polypeptides are highly useful for the immunological detection of autoantibodies

reactive with them since such autoantibodies are indicative of autoimmune rheumatic diseases.

The cDNA clones produced by the present invention encode human topoisomerase I is demonstrated by the fact that: (1) partial proteolysis maps of the cloned fusion protein and naturally-occurring nuclear topoisomerase I are virtually identical; (2) the 67,700 Dalton carboxy-terminal portion of the cloned protein exhibits topoisomerase I activity in vitro; and (3) the cloned cDNA shows a high degree of sequence similarity with topoisomerase I from yeast (Thrash, et al., ibid). The hybridization data obtained indicate that human cells contain only a single locus complementary to this cDNA and that this locus gives rise to a single mRNA.

Topoisomerase I shows a significant degree of evolutionary amino acid sequence conservation (42%) between man and yeast. A comparison of the polypeptide sequences of human and yeast enzymes (Figure 6A) reveals two regions of relatively stronger sequence similarity which are designated region I (435 residues, starting at 197 in the humans) and region II (the carboxy-terminal 70 residues). At optimal alignment both enzymes share a 54% amino acid sequence identity which increases to 69% when conservative substitutions are included. A schematic which indicates the relative locations and extent of regions I and II is shown in Figure 6B.

Interestingly, human and yeast topoisomerase I are predicted to be nearly the same molecular weight, despite the presence of approximately 70 additional residues in the yeast sequence between regions I and II. These additional sequences in the yeast enzyme are balanced by the fact that in the human sequence 70 additional residues are present at the amino terminus.

The amino terminal region of the human protein is remarkable for its sequence composition wherein 65% of the residues are charged based on the fact that of the 197 residues present in the non-similar region between yeast and human, 80 are basic and 48 are acidic. Remarkably, the entire protein has 43% charged residues overall.

The cloned human topoisomerase I was found to have no significant sequence similarities with E. coli type I topoisomerase I.

Although it has been previously shown that topoisomerase I from yeast and E. coli share a small region of sequence similarity, this region was not found in the human sequence for topoisomerase I.

Studies on the topoisomerase I of the invention clearly indicate that the catalytic site is present in the carboxy-terminal 68 kDa region of the polypeptide. The highly charged amino terminal portion of the polypeptide may modulate the activity of the catalytic moiety or, alternatively, be involved in protein-protein interactions. Possible targets for such interactions include regulatory proteins, components of the transcriptional apparatus, or perhaps histones.

In the present invention cDNA clones encoding human topoisomerase I were obtained from a library of human placental mRNA using the lambda gt 11 expression method (Young, et al., Science, 222:778, 1983). Clones expressing human topoisomerase I were screened by immuno-probing with an anti-topoisomerase I autoimmune serum.

The development of specific DNA sequences encoding topoisomerase I polypeptide for splicing into DNA vectors can be accomplished using a variety of techniques. For example, alternative methods which can be employed include: (1) the isolation of a double-stranded DNA sequence from the genomic DNA of the eukaryote; (2) the chemical manufacture of a DNA sequence to provide the necessary codons for the polypeptide of interest; and (3) the in vitro synthesis of a double-stranded DNA sequence by reverse transcription of mRNA isolated from a eukaryotic donor cell. In the latter case, a DNA complement of mRNA is formed which is generally referred to as cDNA.

The manufacture of DNA sequences is frequently the method of choice when the entire sequence of amino acid residues of the desired polypeptide product is known. When the entire sequence of amino acid residues of the desired polypeptide is not known, the direct manufacture of DNA sequences is not possible and the method of choice is the formation of cDNA sequences. Among the standard procedures for isolating cDNA sequences of interest is the formation of plasmid-carrying cDNA libraries which are derived from reverse

transcription of mRNA which is abundant in donor cells which have a high level of genetic expression. In those cases where significant portions of the polypeptide's amino acid sequence are known, the production of labeled single stranded DNA probe sequences duplicating a sequence putatively present in the target cDNA may be employed in DNA/DNA hybridization procedures which are carried out on cloned copies of the cDNA which have been denatured into a single stranded form (Jaye, et al., Nucleic Acid Research, 11:2325, 1983).

Hybridization procedures are useful for the screening of recombinant clones by using labeled mixed synthetic oligonucleotide probes wherein each is potentially the complete complement of a specific DNA sequence in the hybridization sample which includes a heterogeneous mixture of single stranded DNAs or RNAs. These procedures are particularly useful in the detection of cDNA clones derived from sources where extremely low amount of mRNA sequences relating to the polypeptide of interest are present. In other words, by using stringent hybridization conditions directed toward avoidance of non-specific binding it is possible, for example, to allow the autoradiographic visualization of a specific cDNA clone by the hybridization of the target DNA to that single probe in the mixure which is its complete complement (Wallace, et al., Nucleic Acid Research, 9:879, 1981).

Of the three above-noted methods for developing specific DNA sequences for use in recombinant procedures the use of genomic DNA isolates is the least common. This is especially true when it is desirable to obtain the microbial expression of mammalian polypeptides.

The present invention provides novel polypeptides of eukaryotic topoisomerase I which have part or all of the primary structural confirmation, that is, a continuous sequence of amino acid residues, having at least one epitope for autoantibodies to eukaryotic topoisomerase I.

When the eukaryotic topoisomerase I encoded by the DNA sequence of the invention is expressed in a prokaryotic host it is almost immediately degraded into a spectrum of polypeptide fragments. As a consequence, it was surprising the find that these

polypeptide fragments retain at least one epitope capable of reacting with autoantibodies to eukaryotic topoisomerase I. Thus, it is possible to use the polypeptide fragments of the invention rather than intact eukaryotic topoisomerase I to detect autoantibodies to eukaryotic topoisomerase I.

The term "polypeptide," as applied to eukaryotic topoisomerase I polypeptide, denotes any sequence of amino acids having an epitope for autoantibodies to eukaryotic topoisomerase I, wherein the sequence of amino acids is encoded by all or part of the cDNA sequences of the invention.

The polypeptides resulting from microbial expression of the DNA sequences of the invention can be further characterized by their freedom from association with other human polypeptides or other contaminants which might otherwise be associated with eukaryotic topoisomerase I in its natural cellular environment or in such extracellular fluids as plasma or urine.

The term "host" is meant to include not only prokaryotes, but also such eukaryotes as yeast, filamentous fungi, as well as plant and animal cells which can replicate and express an intron-free DNA sequence of eukaryotic topoisomerase I. However, prokaryotes are preferred as the host organism.

The term "prokaryotes" is meant to include all bacteria which can be transformed or transfected with the eukaryotic gene for the expression of eukaryotic topoisomerase I. Procaryotic hosts may include gram negative as well as gram positive bacteria such as, for example, E. coli, S. typhimurium, Serratia marcescens and Bacillus subtilis.

A recombinant DNA molecule coding for the eukaryotic topoisomerase I polypeptides can be used to transform or transfect the host using any of the techniques commonly known to those of ordinary skill in the art. Especially preferred is the use of a plasmid or a virus containing the eukaryotic topoisomerase I coding sequence for purposes of prokaryotic transformation or transfection, respectively.

Methods for preparing fused, operably linked genes and expressing them in bacteria are well-known in the art (Maniatis, et

al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1981). The genetic constructs and methods described therein can be utilized for expression of eukaryotic topoisomerase I in prokaryotic hosts.

In general, expression vectors containing promoter sequences which facilitate the efficient transcription of the inserted eukaryotic genetic sequence are used in connection with the host. The expression vector typically contains an origin of replication, a promoter, and a terminator, as well as specific genes which are capable of providing phenotypic selection of the transformed cells. The transformed prokaryotic hosts can be grown in fermentors and cultured according to techniques known in the art to achieve optimal cell growth. The polypeptides of the invention can then be isolated from the growth medium, cellular lysates, or cellular membrane fractions.

The isolation and purification of the microbially expressed polypeptides of the invention may be by any conventional means such as, for example, preparative chromatographic separations and immunological separations such as those involving the use of monoclonal or polyclonal antibody.

By having provided the sequence of amino acid residues of eukaryotic topoisomerase I, the present invention provides for the manufacture of DNA sequences which code for the host expression of polypeptide analogs or derivatives of eukaryotic topoisomerase I which differ from naturally-occurring forms in terms of the identity or location of one or more amino acid residues and which share some or all of the epitopes of naturally-occurring polypeptide forms.

The novel DNA sequences of the invention include all sequences useful in providing the expression in prokaryotic or eukaryotic host cells of polypeptides which have at least a part of the primary structural confirmation for one or more epitopes capable of reacting with autoantibodies to eukaryotic topoisomerase I which are comprehended by: (a) the DNA sequence as set forth in Figure 5 or its complementary strand; (b) DNA sequences which hybridize to DNA sequences defined in (a) or fragments thereof; and (c) DNA sequences which, but for the degeneracy of the genetic code, would hybridize to

DNA sequences defined in (a) and (b) above. Specifically comprehended in (b) are genomic DNA sequences which encode allelic variant forms of eukaryotic topoisomerase I. Part (c) specifically comprehends the manufacture of DNA sequences which encode eukaryotic topoisomerase I, eukaryotic topoisomerase I fragments and eukaryotic topoisomerase I analogs wherein the DNA sequences thereof may incorporate codons which facilitate translation of mRNA in non-vertebrate hosts.

Since the cDNA sequence of the invention encodes essentially the entire human topoisomerase I molecule it is now a matter of routine to prepare, subclone and express smaller polypeptide fragments of cDNA from this or a corresponding cDNA sequence which would encode as few as one epitope for autoantibodies to human topoisomerase I. The presence of such an epitope on a cloned polypeptide can then be confirmed using, for example, sera from a patient with autoantibodies to topoisomerase I.

The eukaryotic topoisomerase I of the invention is particularly suited for use in immunoassays in which it can be utilized in liquid phase or bound to a solid phase carrier. In addition, the eukaryotic topoisomerase I used in these assays can be detectably labeled in various ways.

Examples of immunoassays which can utilize eukaryotic topoisomerase I of the invention are competitive and non-competitive immunoassays in either a direct or indirect format. Examples of such immunoassays are the radioimmunoassay (RIA) the sandwich (immunometric) assay and the Western blot assay. Detection of antibodies which bind to the eukaryotic topoisomerase I of the invention can be done utilizing immunoassays which run in either the forward, reverse, or simultaneous modes, including immunohistochemical assays on physiological samples. The concentration of eukaryotic topoisomerase I which is used will vary depending on the type of immunoassay and nature of the detectable label which is used. However, regardless of the type of immunoassay which is used, the concentration of eukaryotic topoisomerase I utilized can be readily

determined by one of ordinary skill in the art using routine experimentation.

The eukaryotic topoisomerase I of the invention can be bound to many different carriers and used to detect the presence of antibody specifically reactive with the polypeptide. Examples of well-known carriers include glass, polystyrene, polyvinyl chloride, polypropylene, polyethylene, polycarbonate, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, agaroses and magnetite. The nature of the carrier can be either soluble in insoluble for purposes of the invention. Those skilled in the art will know of other suitable carriers for binding eukaryotic topoisomerase I, or will be able to ascertain such, using routine experimentation.

There are many different labels and methods of labeling known to those of ordinary skill in the art. Examples of the types of labels which can be used in the present invention include enzymes, radioisotopes, colloidal metals, fluorescent compounds, chemiluminescent compounds, and bioluminescent compounds.

For purposes of the invention, the antibody which binds to eukaryotic topoisomerase I of the invention may be present in various biological fluids and tissues. Any sample containing a detectable amount of antibodies to eukaryotic topoisomerase I can be used. Normally, a sample is a liquid such as urine, saliva, cerebrospinal fluid, blood, serum and the like, or a solid or semi-solid such as tissue, feces and the like.

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples which are provided herein for purposes of illusration only and are not intended to limit the scope of the invention.

EXAMPLE 1

CLONING AND EXPRESSION OF TOPOISOMERASE I

A. **Recombinant DNA Procedures.** cDNA clones were obtained using the lambda gt11 expression method of Young and Davis (Science, 222:778, 1983). $1x10^6$ recombinant phage (lambda gt11 library of human placental polyadenosine mRNA; Clontech) were

plated and screened by immunoprobing with an autoimmune anti-topoisomerase I serum (Shero, et al., Science, 231:737, 1986), using standard methods (Maniatis, et al., ibid). DNA and RNA hybridizations were performed according to standard procedures (Maniatis, et al., ibid). Blots were hybridized at 42°C in the presence of 50% formamide and washed at 50°C with a solution of 15 mM NaCl, 1.5 mM Na citrate and 0.1% $NaDodSo_4$ (Maniatis, et al., ibid). cDNA sequences were determined by constructing an exonuclease III deletion series of clone T1A as described by Henicoff (Gene, 28:351, 1984). The deleted fragments were subcloned into M13 for sequencing by the dideoxynucleotide chain termination method (Sanger, et al., Proceedings of the National Academy of Sciences, U.S.A., 74:5463, 1977). Other sequences were determined by cloning specific restriction fragments into M13.

B. Expression of Topoisomerase I as a Fusion Protein in Bacteria. The cloned cDNA T1A (see Figure 1) inserted into the pATH11 expression vector (provided by T.J. Koerner, Duke University) was expressed in E. coli (DH1) as a fusion protein comprised of 32kDa of the bacterial trpE protein linked to 67.7 kDa of the protein encoded by the cDNA, as previously described (Earnshaw, et al, Journal of Cell Biology, 104, 817, 1987). This fusion protein (fp-topo I) was induced and visualized by immunoblotting as described (Earnshaw, et al., ibid). All immunoblotting was done with the same anti-topoisomerase I autoimmune serum.

1×10^6 recombinant phage were screened with topoisomerase I antiserum (Shero, et al., ibid). Eight immunoreactive plaques were purified and the longest cDNA clone (T1A, 2929 bp) was analyzed by restriction mapping and sequenced according to the strategy outlined in Figure 1. Longer cDNAs were obtained by screening additional recombinant phage with the 5' Eco RI - Nde I fragment of T1A as a hybridization probe. 10 additional clones were obtained including clone T1B (3645 bp, Figure 1), which appears to contain the entire coding region (see below). Both T1A and T1B terminate at their 5' ends with an Eco RI site (located in the 3' nontranslated region) which

was also found in several other clones obtained from both the antibody and DNA hybridization screens.

Clone T1A was subcloned into a bacterial expression plasmid (pATH11, in Earnshaw, et al., ibid) in which protein encoded by the cDNA is linked in a fusion protein to the 32 kDa amino terminal portion of the bacterial trpE protein. Bacterial lysates harboring pATH11 containing T1A (pATH11-T1A) contained several immunoreactive polypeptides (probably proteolysis fragments) that were not evident in lysates harboring only pATH11 (Figure 2A).

EXAMPLE 2

CHARACTERIZATION OF CLONED TOPOISOMERASE I

A. **Proteolysis Mapping.** Proteolysis maps of human topoisomerase I and fp-topo I were performed as described (Cleveland, Methods in Enzymology, 96:222, 1983), except that immunoblotting was used to visualize the proteolysis fragments instead of Coomassie blue staining. Gel bands corresponding to human topoisomerase I and fp-topo I (the largest two bands) were excised in small strips, inserted into the wells of another $NaDodSO_4$ gel, and overlaid with either chymotrypsin (100 g) or V8 protease from Staphylococcus aureus (1.0 or 10.0 g). Electrophoresis was initiated at 50 volts until the dye completely entered the stacker. Proteolysis occurs as the proteins stack (Cleveland, ibid).

To examine the biochemical similarity between the fusion protein (fp-topo I) and human nuclear topoisomerase I, their proteolysis maps were compared. HeLa nuclear topoisomerase I and fp-topo I (the largest two species seen in Figure 2A) were obtained by excision from $NaDodSO_4$ polyacrylamide gels. Human nuclear topoisomerase I and fp-topo I gave rise to nearly identical patterns of immunoreactive proteolytic fragments, indicating that the two polypeptides are closely related (Figure 2B). The virtual identity of these proteolysis maps indicates that the immunoreactive proteolysis fragments of fp-topo I contain nondetectable amounts of trpE.

B. **DNA Topoisomerase I Assays.** Assays for bacterial and human topoisomerase I were performed essentially as described by Thrash et al. (ibid).

To confirm that clone T1A encoded topoisomerase I, fp-topo I was assayed for Mg^{++}-independent supercoil-relaxing activity, a characteristic of eukaryotic topoisomerase I (Thrash, et al., ibid). Bacteria expressing fp-topo I (lanes +), and nonexpressing control bacteria (harboring pATH11 only, lanes –) were lysed, and after pelleting debris, aliquots of the supernatants were tested for their ability to relax super coiled DNA (Figure 3A). In buffer P which contains excess Mg^{++}, lysates of bacteria that were not expressing fp-topo I contained supercoil-relaxing activity due to bacterial topoisomerase I, which requires Mg^{++} as a cofactor (-, lane P). In buffer E, which does not contain free Mg^{++}, bacterial topoisomerase I activity is abolished (-,lane E). In lysates of bacteria expressing fp-topo I, supercoil-relaxing activity is evident both in the presence and absence of free Mg^{++} (+, lanes P and E). This Mg^{++}-independent activity, characteristic of eukaryotic type I topoisomerases, must be due to fp-topo I. The difference in the distribution of topoisomers produced by bacterial and human type I topoisomerases is characteristic of the enzymes (Thrash, et al., ibid).

To demonstrate further that fp-topo I encodes human topoisomerase I, lysates of bacteria expressing fp-topo I were immunoprecipitated with anti-topoisomerase I autoimmune serum. Mg^{++}-independent topoisomerase I activity was removed from these lysates, but not from control lysates treated with normal human serum (Figure 3B). This data shows that T1A encodes a catalytically active portion of human topoisomerase I.

To investigate whether topoisomerase I is encoded by a single gene or a family of genes as has been seen for many other eukaryotic proteins such as tubulins (Cleveland, et al., Annual Review of Biochemistry, 54:331, 1985), blots of human genomic DNA that had been digested with either Eco R1 or Hind III were analyzed. A DNA segment spanning codons 484 to 578 (fragment with asterisk in Figure 1, and see Figure 6A) was prepared from T1A and used as a hybridization probe. The resultant autoradiograph is shown in Figure 4A. Only a single DNA fragment was detected in either digest. Since moderate conditions of stringency were utilized, only a single locus highly

homologous to the cloned topoisomerase I sequence is apparently present in the human genome.

To determine the size of the mRNA that encodes human topoisomerase I, we probed blots of HeLa cell poly(A) RNA with the entirety of T1A. A single RNA species (Figure 4B) of 4100 nucleotides was detected.

C. The Deduced Amino Acid Sequence of Human Topoisomerase I.

The nucleotide sequence of the human topoisomerase I cDNA contained in clones T1A and T1B was determined according to the strategy shown in Figure 1. The 3645 base T1B sequence begins with a putative 5' untranslated region of 211 nucleotides, followed by a 765 amino acid open reading frame, and terminates with 1139 nucleotides of 3' untranslated sequence. The encoded polypeptide has a molecular weight of 90,649 Da, and a calculated isoelectric point of 10.05 (26% basic and 18% acidic residues). The deduced molecular weight is somewhat smaller than the 100,000 Da estimate for authentic topoisomerase I determined by $NaDodSO_4$-PAGE(Liu, et al, Proceedings of the National Academy of Sciences, U.S.A., 78:3487, 1981), however, apparent molecular weights may differ substantially from actual molecular weights (Earnshaw, et al, ibid).

Since the corresponding mRNA is 4100 bases in length, it is estimated that clone T1B is 250-350 amino acids short of full length assuming a poly(A) tail of 100-200. The missing sequence may be either 5' or 3' to the open reading frame. Since the 3' end of the clones is bounded by a natural Eco RI site, the true length of the 3' nontranslated region is unknown.

It is not known whether the cDNA sequence presented in Figure 5 represents the entire protein coding region. Although the unambiguous identification of initiation methionines from cDNA sequences is problematic, the specified initiator codon appears to be correct for three reasons. 1) The first AUG in the clone is in frame with a translation unit that encodes a protein of approximately the same size as authentic human topoisomerase I. 2) This AUG codon is immediately preceded by nucleotides that have been associated with

eukaryotic start codons (Kozak, Cell, 44:283, 1986): a purine is in position -3, and cytosines are in positons -1, -4 and -5. 3) The 211 bp sequence upstream of the putative start codon is 73% GC rich and is not deficient in CpG (the 211 bp sequence downstream of the putative start codon is only 40% GC rich). Such a sequence arrangement is characteristic of the 5' ends of many genes that do not display tissue specific expression (Bird, Nature, 321:209, 1986). Despite this, it should be noted that an open reading frame extends from the presumptive initiation codon all the way to the 5' end of the clone.

EXAMPLE 3
DETECTION OF AUTOANTIBODIES TO TOPOISOMERASE I IN CLINICAL SPECIMENS

Sera from normal individuals and from thirteen patients suspected of having autoantibodies to human topoisomerase I were tested using the Western blot technique (Earnshaw, et al., Journal of Clinical Investigation, 72:425, 1986) and the polypeptides expressed by the cDNA sequence of the invention. The thirteen sera from patients suspected of having autoantibodies to human topoisomerase I had previously shown a positive reaction for the presence of such autoantiodies when tested in a commercially available diagnostic test which utilizes a crude extract from mammalian nuclei. When these same thirteen sera were tested against the polypeptides of the invention only eleven of the patient's sera were positive. Since the Western technique is more sensitive than the technique utilized in the commercial test, two of these patients may have been false positive. Since the commercial test contains other nuclear antigens in addition to topoisomerase I such as, for example, the centromere, these patients sere may have reacted to one or more of these other non-topoisomerase I antigens in the commercial test. None of the normal sera which were tested reacted with the polypeptides of the invention.

The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing form the spirit or scope of the invention.

CLAIMS

1. A cDNA sequence consisting essentially of a cDNA sequence encoding human topoisomerase I.

2. A host cell transformed or transfected with the cDNA sequence of claim 1.

3. A biologically functional plasmid or viral DNA vector including the cDNA sequence of claim 1.

4. A host cell stably transformed or transfected with a DNA vector according to claim 3.

5. A cDNA sequence consisting essentially of a cDNA sequence encoding a polypeptide having an amino acid sequence possessing at least one epitope for autoantibodies to eukaryotic topoisomerase I.

6. The cDNA sequence according to claim 5 and including the protein coding region set forth in Figure 5.

7. The cDNA sequence of claim 5, wherein said cDNA sequence codes for the expression of human topoisomerase I.

8. A host cell transformed or transfected with a the cDNA sequence of claim 5.

9. A method of providing eukaryotic topoisomerase I polypeptide having at least one epitope for autoantibodies to eukaryotic topoisomerase I comprising:

(a) providing a host cell which replicates and expresses an intron-free DNA sequence of eukaryotic topoisomerase I polypeptide;

(b) growing said host cell; and

(c) recovering said eukaryotic topoisomerase I polypeptide.

10. The method of claim 9, wherein said eukaryotic topoisomerase I is human topoisomerase I.

11. A method of detecting autoantibodies to eukaryotic topoisomerase I which comprises:

(a) contacting a sample with a polypeptide encoded by the cDNA sequence of claim 1;

(b) incubating the components of step (a) for a period of time and under conditions sufficient for said autoantibodies to bind to said polypeptide;

(c) separating said autoantibodies bound to said polypeptide from said sample; and

(d) detecting the presence of said autoantibodies bound to said polypeptide.

12. The method of claim 11 wherein said eukaryotic topoisomerase I is human topoisomerase I.

13. The method of claim 11 wherein said sample is from a human.

14. The method of claim 11 wherein said detecting utilizes a detectably labeled protein capable of binding to said autoantibody.

15. The method of claim 14 wherein said binding protein is a detectably labeled second antibody.

16. The method of claim 15 wherein said detectable label is selected from the group consisting of a radioisotope, a fluorescent compound, a colloidal metal, a chemiluminescent compound, a bioluminescent compound and an enzyme.

FIG. 1

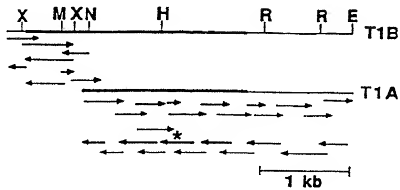
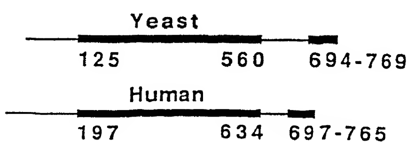

X
M
X
N
H
R
R
E
T1B
T1A
*
1 kb

FIG. 6B


Yeast
125
560
694-769
Human
197
634
697-765

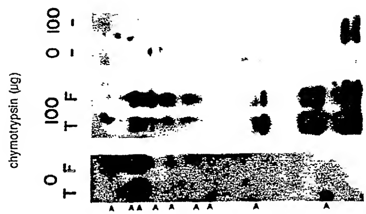


FIG.2C

FIG.2B

FIG.2A

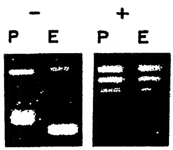


FIG.3A

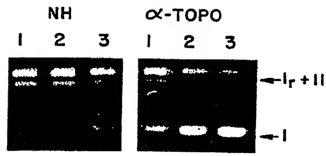


FIG.3B

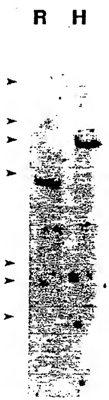


FIG.4A

FIG.4B

FIG. 5A

GAATTCGGCGGCCGCCCGCCCGGCAGTCAGGCAGCGTCGGCGCCGTGGTAGCAGCCTCAGCCGTTTCTGGAGTCTCGGGCCCAC
AGTCACCGCCGCTTACCTGCGCCTCCTCGAGCCTCCGGAGTCCCCGTCCGCCCGCACAGGCCGGTTCGCCGTCTGCGTCTCCCCCACGCCGCCTCGCCTGCCGCCGCGCTCGTCCCTCCGGGCCGAC

1 32
ATG AGT GGG GAC CAC CTC CAC AAC GAT TCC CAG ATC GAA GCG GAT TTC CGA TTG AAT GAT TCT CAT AAA CAC AAA GAT AAA CAC AAA GAT CGA GAA
MET Ser Gly Asp His Leu His Asn Asp Ser Gln Ile Glu Ala Asp Phe Arg Leu Asn Asp Ser His Lys His Lys Asp Lys His Lys Asp Arg Glu

33 64
CAC CGG CAC AAA GAA CAC AAG AAG GAG AAG GAC CGG GAA AAG TCC AAG CAT AGC AAC AGT GAA CAT AAA GAT TCT GAA AAG AAA CAC AAA GAG AAG
His Arg His Lys Glu His Lys Lys Glu Lys Asp Arg Glu Lys Ser Lys His Ser Asn Ser Glu His Lys Asp Ser Glu Lys Lys His Lys Glu Lys

65 96
GAG AAG ACC AAA CAC AAA GAT GGA AGC TCA GAA AAG CAT AAA GAC AAA CAT AAA GAC AGA GAC AAG GAA AAA CGA AAA GAG GAA AAG GTT CGA GCC
Glu Lys Thr Lys His Lys Asp Gly Ser Ser Glu Lys His Lys Asp Lys His Lys Asp Arg Asp Lys Glu Lys Arg Lys Glu Glu Lys Val Arg Ala

97 128
TCT GGG GAT GCA AAA ATA AAG AAG GAG AAG GAA AAT GGC TTC TCT AGT CCA CCA CAA ATT AAA GAT GAA CCT GAA GAT GAT GGC TAT TTT GTT CCT
Ser Gly Asp Ala Lys Ile Lys Lys Glu Lys Glu Asn Gly Phe Ser Ser Pro Pro Gln Ile Lys Asp Glu Pro Glu Asp Asp Gly Tyr Phe Val Pro

129 160
CCT AAA GAG GAT ATA AAG CCA TTA AAG AGA CCT CGA GAT GAG GAT GAT GTT GAT TAT AAA CCT AAG AAA ATT AAA ACA GAA GAT ACC AAG AAG GAG
Pro Lys Glu Asp Ile Lys Pro Leu Lys Arg Pro Arg Asp Glu Asp Asp Val Asp Tyr Lys Pro Lys Lys Ile Lys Thr Glu Asp Thr Lys Lys Glu

161 192
AAG AAA AGA AAA CTA GAA GAA GAA GAG GAT GGT AAA TTG AAA AAA CCC AAG AAT AAA GAT AAA GAT AAA AAA GTT CCT GAG CCA GAT AAC AAG AAA
Lys Lys Arg Lys Leu Glu Glu Glu Glu Asp Gly Lys Leu Lys Lys Pro Lys Asn Lys Asp Lys Asp Lys Lys Val Pro Glu Pro Asp Asn Lys Lys

193 224
AAG AAG CCG AAG AAA GAA GAG GAA CAG AAG TGG AAA TGG TGG GAA GAA GAG CGC TAT CCT GAA GGC ATC AAG TGG AAA TTC CTA GAA CAT AAA GGT
Lys Lys Pro Lys Lys Glu Glu Glu Gln Lys Trp Lys Trp Trp Glu Glu Glu Arg Tyr Pro Glu Gly Ile Lys Trp Lys Phe Leu Glu His Lys Gly

225 256
CCA GTA TTT GCC CCA CCA TAT GAG CCT CTT CCA GAG AAT GTC AAG TTT TAT TAT GAT GGT AAA GTC ATG AAG CTG AGC CCC AAA GCA GAG GAA GTA
Pro Val Phe Ala Pro Pro Tyr Glu Pro Leu Pro Glu Asn Val Lys Phe Tyr Tyr Asp Gly Lys Val MET Lys Leu Ser Pro Lys Ala Glu Glu Val

257 288
GCT ACG TTC TTT GCA AAA ATG CTC GAC CAT GAA TAT ACT ACC AAG GAA ATA TTT AGG AAA AAT TTC TTT AAA GAC TGG AGA AAG GAA ATG ACT AAT
Ala Thr Phe Phe Ala Lys MET Leu Asp His Glu Tyr Thr Thr Lys Glu Ile Phe Arg Lys Asn Phe Phe Lys Asp Trp Arg Lys Glu MET Thr Asn

289 320
GAA GAG AAG AAT ATT ATC ACC AAC CTA AGC AAA TGT GAT TTT ACC CAG ATG AGC CAG TAT TTC AAA GCC CAG ACG GAA GCT CGG AAA CAG ATG AGC
Glu Glu Lys Asn Ile Ile Thr Asn Leu Ser Lys Cys Asp Phe Thr Gln MET Ser Gln Tyr Phe Lys Ala Gln Thr Glu Ala Arg Lys Gln MET Ser

FIG. 5B

321
AAG GAA GAG AAA CTG AAA ATC AAA GAG GAG AAT GAA AAA TTA CTG AAA GAA TAT GGA TTC TGT ATT ATG GAT AAC CAC AAA GAG AGG ATT GCT AAC
Lys Glu Glu Lys Leu Lys Ile Lys Glu Glu Asn Glu Lys Leu Leu Lys Glu Tyr Gly Phe Cys Ile MET Asp Asn His Lys Glu Arg Ile Ala Asn 352
353
TTC AAG ATA GAG CCT CCT GGA CTT TTC CGT GGC CGC GGC AAC CAC CCC AAG ATG GGC ATG CTG AAG AGA CGA ATC ATG CCC GAG GAT ATA ATC ATC
Phe Lys Ile Glu Pro Pro Gly Leu Phe Arg Gly Arg Gly Asn His Pro Lys MET Gly MET Leu Lys Arg Arg Ile MET Pro Glu Asp Ile Ile Ile 384
385
AAC TGT AGC AAA GAT GCC AAG GTT CCT TCT CCT CCT CCA GGA CAT AAG TGG AAA GAA GTC CGG CAT GAT AAC AAG GTT ACT TGG CTG GTT TCC TGG
Asn Cys Ser Lys Asp Ala Lys Val Pro Ser Pro Pro Pro Gly His Lys Trp Lys Glu Val Arg His Asp Asn Lys Val Thr Trp Leu Val Ser Trp 416
417
ACA GAG AAC ATC CAA GGT TCC ATT AAA TAC ATC ATG CTT AAC CCT AGT TCA CGA ATC AAG GGT GAG AAG GAC TGG CAG AAA TAC GAG ACT GCT CGG
Thr Glu Asn Ile Gln Gly Ser Ile Lys Tyr Ile MET Leu Asn Pro Ser Ser Arg Ile Lys Gly Glu Lys Asp Trp Gln Lys Tyr Glu Thr Ala Arg 448
449
CGG CTG AAA AAA TGT GTG GAC AAG ATC CGG AAC CAG TAT CGA GAA GAC TGG AAG TCC AAA GAG ATG AAA GTC CGG CAG AGA GCT GTA GCC CTG TAC
Arg Leu Lys Lys Cys Val Asp Lys Ile Arg Asn Gln Tyr Arg Glu Asp Trp Lys Ser Lys Glu MET Lys Val Arg Gln Arg Ala Val Ala Leu Tyr 480
481
TTC ATC GAC AAG CTT GCT CTG AGA GCA GGC AAT GAA AAG GAG GAA GGA GAA ACA GCG GAC ACT GTG GGC TGC TGC TCA CTT CGT GTG GAG CAC ATC
Phe Ile Asp Lys Leu Ala Leu Arg Ala Gly Asn Glu Lys Glu Glu Gly Glu Thr Ala Asp Thr Val Gly Cys Cys Ser Leu Arg Val Glu His Ile 512
513
AAT CTA CAC CCA GAG TTG GAT GGT CAG GAA TAT GTG GTA GAG TTT GAC TTC CTC GGG AAG GAC TCC ATC AGA TAC TAT AAC AAG GTC CCT GTT GAG
Asn Leu His Pro Glu Leu Asp Gly Gln Glu Tyr Val Val Glu Phe Asp Phe Leu Gly Lys Asp Ser Ile Arg Tyr Tyr Asn Lys Val Pro Val Glu 544
545
AAA CGA GTT TTT AAG AAC CTA CAA CTA TTT ATG GAG AAC AAG CAG CCC GAG GAT GAT CTT TTT GAT AGA CTC AAT ACT GGT ATT CTG AAT AAG CAT
Lys Arg Val Phe Lys Asn Leu Gln Leu Phe MET Glu Asn Lys Gln Pro Glu Asp Asp Leu Phe Asp Arg Leu Asn Thr Gly Ile Leu Asn Lys His 576

FIG. 5C

577 608

CTT CAG GAT CTC ATG GAG GGC TTG ACA GCC AAG GTA TTC CGT ACG TAC AAT GCC TCC ATC ACG CTA CAG CAG CAG CTA AAA GAA CTG ACA GCC CCG

Leu Gln Asp Leu MET Glu Gly Leu Thr Ala Lys Val Phe Arg Thr Tyr Asn Ala Ser Ile Thr Leu Gln Gln Gln Leu Lys Glu Leu Thr Ala Pro

609 640

GAT GAG AAC ATC CCA GCG AAG ATC CTT TCT TAT AAC CGT GCC AAT CGA GCT GTT GCA ATT CTT TGT AAC CAT CAG AGG GCA CCA CCA AAA ACT TTT

Asp Glu Asn Ile Pro Ala Lys Ile Leu Ser Tyr Asn Arg Ala Asn Arg Ala Val Ala Ile Leu Cys Asn His Gln Arg Ala Pro Pro Lys Thr Phe

641 672

GAG AAG TCT ATG ATG AAC TTG CAA ACT AAG ATT GAT GCC AAG AAG GAA CAG CTA GCA GAT GCC CGG AGA GAC CTG AAA AGT GCT AAG GCT GAT GCC

Glu Lys Ser MET MET Asn Leu Gln Thr Lys Ile Asp Ala Lys Lys Glu Gln Leu Ala Asp Ala Arg Arg Asp Leu Lys Ser Ala Lys Ala Asp Ala

673 704

AAG GTC ATG AAG GAT GCA AAG ACG AAG AAG GTA GTA GAG TCA AAG AAG AAG GCT GTT CAG AGA CTG GAG GAA CAG TTG ATG AAG CTG GAA GTT CAA

Lys Val MET Lys Asp Ala Lys Thr Lys Lys Val Val Glu Ser Lys Lys Lys Ala Val Gln Arg Leu Glu Glu Gln Leu MET Lys Leu Glu Val Gln

705 736

GCC ACA GAC CGA GAG GAA AAT AAA CAG ATT GCC CTG GGA ACC TCC AAA CTC AAT TAT CTG GAC CCT AGG ATC ACA GTG GCT TGG TGC AAG AAG TGG

Ala Thr Asp Arg Glu Glu Asn Lys Gln Ile Ala Leu Gly Thr Ser Lys Leu Asn Tyr Leu Asp Pro Arg Ile Thr Val Ala Trp Cys Lys Lys Trp

737 765

GGT GTC CCA ATT GAG AAG ATT TAC AAC AAA ACC CAG CGG GAG AAG TTT GCC TGG GCC ATT GAC ATG GCT GAT GAA GAC TAT GAG TTT TAG CCAGTCT

Gly Val Pro Ile Glu Lys Ile Tyr Asn Lys Thr Gln Arg Glu Lys Phe Ala Trp Ala Ile Asp Met Ala Asp Glu Asp Tyr Glu Phe *

CAAGAGCCAGAGTTCTGTGAAGAGGAACAGTGTGGTTTGGGAAAGATGGATAAACTGAGCCTCACTTGCCCTCGTGCCTGGGGGAGAGAGGCAGCAAGTCTTAACAAACCAACATCTTTGCGAAAAG

ATAAACCTGGACATATTATATAAGGGAGACCTGAGCCAGTTGTCCTATGGACAACTTATTTAAAAATATTTCAGATATCAAAATTCTAGCTGTATGATTTGTTTTGAATTTTGTTTTTATTTTCAAGAG

CCCAAGTGGATCGGAATTTGTCAGCGTTCTACCAGGCAAATTCACTGTTTCACTGAAATGTTTGGATTCTCTTAGCTACTGTATGCAAAGTCCGATTATATTGGTGCGTTTTTACAGTTAGGGTTTT

CCAATAACTTCTATATTTTAATAGAAATAAATTCCTAAACTCCCTTCCCTCTCTCCCATTTCACGAATTTAAAATTAAGTAGAACAAAAAACCCAGCGCACCTGTTAGAGTCGTCACTCTCTATTGT

CATCGGATCAATTTTCATTAAACTTGAAGCAGTCGTGCCTTTCGCAGTGTTTTGGTTCAGACACCTGTTCACAGAAAAAGCATGATGGGAAAATATTTCCTGACTTGAGTGTTCCTTTTTAAATGT

GAATTTTTTTTTTTTTTAATTATTTTAAAATATTTAAACCTTTTTCTTGATCTTAAAGATCGTGTAGATTGGGGTTGGGGAGGGATGAAGGGCGAGTGAATCTAACGATAATGAAATAATCAGTGAC

TGAAACCATTTTCCCATCATCCTTTGTTCTGAGCATTCGCTGTACCCTTTAAGATATCCATCTTTTTCTTTTTAACCCTAATCTTTCACTTGAAAGATTTTATTGTATAAAAAGTTTCACAGGTCAA

TAAACTTAGAGCAAAATGAGTATTTGGTCCAAAAAAAGGAAAAATAATCAAGATTTTAGGCTTTTATTTTTTCTTTTGTAATTGTGTAAAAAATGGAAAAAAACATAAAAAGCACAATTTTAATGT

CAAGACATTTTTTCCTATAATCATTAGTTTTAGAGGCATTGTTAGTTTAGTGTGTGTGCAGAGTCCATTTCCCACATCTTTCCTCAAGTATCTTCTATTTTTATCATGAATTC

FIG. 6A

```
                                                                                                    1
                                                                                                    MTIADASK
                                          MSGDHLHNDSQIEADFRLNDSHKHKDKHKDREHRHKEHKKEKDREKSKHSNSEHKDSEKKHKEKEKTKHKDGSSEKHKDKHKD
                                          1         10        20        30        40        50        60        70        80
12        22        32        42        52        62        72        82        92        102       112       122       132       142
VNHELSSDDDDDVPLSQTLKKRKVASMNSASLQDEAEPYDSDEAISKISKKKTKKIKTEPVQSSSLPSPPAKKSATSKPKKIKKEDGDVKVKTTKKEEQENEKKKREEEEEEDKKAKEEEEEYKWEKENEDDTIKWVTL
 .:   ..   .   .:: :  ...:.  :  :: . :  .   :   : .:                 :   :.  .:  : .: . .   :  :    . .::  : .    :  ::::   :::: :  .. ::: :
RDKEKRKEEKVRASGDAKIKKEKENGFSSPP QIKDEP EDDGYFVP PKEDIKPLKRPRDEDDADYKPKKIKTEDTKKEKKRKLEEEEDGKLKKPKNKDKDKKVPEPDNKKKKPKKEEEQKWKWWEEERYPEGIKWKFL
 87        97        107       117       127       137       147  154       164       174       184       194       204       214
152       162       172       182       192       202       212       222       232       242       252       262       272       282
KHNGVIFPPPYQPLPSHIKLYYDGKPVDLPPQAEEVAGFFAALLESDHAKNPVFQKNFFNDFLQVLKESGGPINGIEIKEFSRCDFTKMFDYFQLQKEQKKQLTSQEKKQIRLEREKFEEDYKFCEIDGRREQVGNFKVE
 : :..:.::: :::  .: :::::.. : : :::::.:::  :. .     .: :::: :      :: ..   .: :  .:.:::: :  :: .: : .::..  ::  :. : ::   .: :: .:...:   ::::.:
EHKGPVFAPPYEPLPENVKFYYDGKVMKLSPKAEEVATFFAKMLDHEYTTKEIFRKNFFKDW R  KEMINEEKNI ITNLSKCDFTQMSQYFKAQTEARKQMSKEEKLKIKEENEKLLKEYGFCIMDNHKERIANFKIE
 224       234       244       254       264       274       284  291        301       311       321       331       341       351
292       302       312       322       332       342       352       362       372       382       392       402       412       422
PPDLFXGRGAHPKTGKLKRRVNPEDIVLNLSKDAPVPPAPEGHKWGEIRHDNTVQWLAMWRENIFNSFKYVRLAANSSLKGQSDYKKFEKARQLKSYIDAIRRDYTRNLKSKVMLERQKAVAIYLIDVFALRAGGEKSED
:: :: ::: ::: : ::::. ::::..: :::: :: .: :::: :.:::: :.::. : :::  :.::. : ..: .::  :  : : :: :: ..: ::  :     ::: :  ::.:::.:.:: .:::::.:: :
PPGLFRGRGNHPKMGMLKRRIMPEDIIINCSKDAKVPSPPPGHKWKEVRHDNKVIWLVSWTENIQGSIKYIMLNPSSRIKGEKDWQKYETARRLKKCVDKIRNQYREDWKSKEMKVRQRAVALYFIDKLALRAGNEKEEG
 360       370       380       390       400       410       420       430       440       450       460       470       480       490
 432       442        452       462       472       482       492 496       506       516       526       536       546       556
E ADTVGCCSLRYEHVTLKPP NT   VI FDFLGKDSIRFYQEVEVDKQVFKNLTIFKRPPKQPGHQLFDRLDPSILNKYLQNYMPGLTAKVFRTYNASKTMQDQL DLIPNKGSVAEKILKYNAANRTVAILCNHQRT
: ::::::::: ::. :.:     .   :  ::::::::: :. : :.: :::::. :      :::   :::::  .:::: ::   : ::::::::::::: :.: :: .: .    ... ::: :: ::: :::::::::
ETADTVGCCSLRVEHINLHPELDGQEYVVEFDFLGKDSIRYYNKVPVEKRVFKNLQLFMEN KQPEDDLFDRLNTGILNKHLQDLMEGLTAKVFRTYNASITLQQQLKELTAPDENIPAKILSYNRANRAVAILCNHQRA
 500       510       520       530       540       550       560       569       579       589       599       609       619       629
 565       575       585       595        605       615      623       633       643       653       663  670       680       690
VTKGHAQTVEKANNRIQELEWQKIRCKRAILQLDKDL LKKEPKYFEEIDDLTKEDEATIHKRIIDREIEKYQRKFVRENDKRKFEKEELLPESQLKEWLEKVDEKKQEFEKELKTGEVELKSSWNSVEKIKAQVEKLEQ
  :.    ..   ....:      : .  .: . .. :.   :..:.   . . :      .                                                                                 :. :::
PPKTFEKSMMNLQTKIDAKKEQLADARRDLKSAKADAKVMKDAKT KKVVESKKKAVQRL*****************************************************************************EEQLMKLE
 639       649       659       669       679       689                                                                                           698
 705       715       725       735       745       755       765
RIQTSSIQLKDKEENSQVSLGTSKINYIDPRLSVVFCKKYDVPIEKIFTKTLREKFKWAIESVDENWRFX
  : .      : ::: :. :::::.::.:::..:..:::  :::::: .:: :::: :::. .::   :
 VQAT      DREENKQIALGTSKLNYLDPRITVAWCKKWGVPIEKIYNKTQREKFAWAIDMADEDYEFX
            708       718       728       738       748       758
```

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US89/01116

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(4): C07H 21/04; C12N 15/00, 9/90, 1/20; G01N 33/564
US CL: 536/27; 435/7, 172.3, 233, 252.3, 320

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
U.S.	435/6, 7, 172.3, 233, 235, 252.3, 320 536/27; 436/508 935/12, 14, 18, 27, 31, 76, 81, 82

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

Computer Search: Medline, WPI, WPIL, CA Search, APS, IG Suite (GenBank, EMBL, PIR)

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
P,X Y	Proceedings of the National Academy of Sciences, USA, Volume 85, issued April 1988 P. D'Arpa, "cDNA Cloning of Human DNA Topoisomerase I: Catalytic Activity of a 67.7-kDa Carboxy-Terminal Fragment", see page 2543, column 2; sentence bridging columns 1 and 2 on page 2547; Figure 5, page 2545.	1-10 11-16
Y	Chromosoma, Volume 94, No. 2, issued 1986, H. Guldner et al, "Scl 70 Autoantibodies from Scleroderma patients Recognize a 95 kDa Protein Identified as DNA Topoisomerase I", see page 133, column 1; page 134, column 2; page 136, column 2, and page 137, columns 1 and 2.	1-16
Y	FEBS Letters, Volume 209, No. 2, issued 15 December 1986, D. Samuels et al, "Inhibition of Topoisomerase I By Antibodies In Sera From Scleroderma Patients", see page 231, column 1.	1-16

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
23 June 1989	12 JUL 1989
International Searching Authority	**Signature of Authorized Officer**
ISA/US	ESTHER M. KEPPLINGER

Form PCT/ISA/210 (second sheet) (Rev.11-87)

International Application No. PCT/US89/01116

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
Y	Science, Volume 231, issued 14 February 1986, J. H. Shero et al, "High Titers of Autoantibodies To Topoisomerase I (Scl-70) In Sera From Scleroderma Patients", see page 738, right column; page 739, paragraph bridging columns 2 and 3, and page 739, column 3.	1-16
Y	Journal of Immunological Methods, Volume 80, No. 2, issued 25 June 1985, C. Geisler et al, "An Enzyme-Linked Immunosorbent Assay For Autoantibodies Against Nuclear Protein Scl-70", see paragraph bridging pages 213 and 214 and page 219.	11-16
Y	Proceedings of the National Academy of Sciences, USA, Volume 82, issued November 1985, T. Goto et al, "Cloning of Yeast TOP1, The Gene Encoding DNA Topoisomerase I, And Construction of Mutants Defective In Both DNA Topoisomerase I and DNA Topoisomerase II" see page 7178, column 2.	1-16
Y	Proceedings of the National Academy of Sciences, USA, Volume 83, issued July 1986, G. Maul et al, "Topoisomerase I Identified By Scleroderma 70 Antisera: Enrichment of Topoisomerase I At the Centromere In Mouse Mitotic Cells Before Anaphase", see page 5145 column 2 and page 5147, column 1.	1-16